April 1993

GENETIC VARIATION IN DNA OF COHO SALMON FROM THE LOWER COLUMBIA RIVER FINAL REPORT

DOE/BP-30198-1

DOE/BP-30198-1

This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views of this report are the author's and do r necessarily represent the views of BPA.

This document should be cited as follows:

Fobes,Stephen; Kathy Knudsen, Fred Allendorf Division of Biological Sciences University of Montana, U. S Department of Energy, Bonneville Power Administration, Division of Fish and Wildlife, Project Number 92-03, Contract Number DE-BI79-92BP30 198, 29 electronic pages (BPA Report DOE/BP-30198-1)

This report and other BPA Fish and Wildlife Publications are available on the Internet at:

http://www.efw.bpa.gov/cgi-bin/efw/FW/publications.cgi

For other information on electronic documents or other printed media, contact or write to:

Bonneville Power Administration
Environment, Fish and Wildlife Division
P.O. Box 3621
905 N.E. 11th Avenue
Portland, OR 97208-3621

GENETIC VARIATION IN DNA OF COHO SALMON FROM THE LOWER COLUMBIA RIVER

FINAL REPORT

Prepared by:

Stephen Fobes
Kathy Knudsen
Fred Allendorf
Division of Biological Sciences
University of Montana
Missoula, MT 598 12

Prepared for:

U. S Department of Energy
Bonneville Power Administration
Division of Fish and Wildlife
P.O. Box 3621
Portland, OR 97283-362 1

Project Number 92-03 5
Contract Number DE-BI79-92BP30 198

APRIL 1993

ACKNOWLEDGEMENTS

This research was funded by the Bonneville Power Administration (Project No. 92-035), Debbie Watkins, Project Manager. We thank Kathryn Kostow, Oregon Department of Fish and Wildlife (ODFW), Orlay Johnson, National Marine Fisheries Service, and Ken Currens, Oregon Cooperative Fishery Research Unit for providing samples of coho salmon, Tim Schamber (ODFW) for the coho salmon used in experimental matings, and Joan Strange and Tom North of the University of Montana Molecular Biology Facility for automated DNA sequencing services.

ABSTRACT

The goal of this project was to develop techniques to provide the information needed to determine if Lower Columbia River coho salmon represent a "species" under the Endangered Species Act. Our report features two new nuclear DNA approaches to the improved detection of genetic variation: 1) Studies of DNA-level genetic variation for two nuclear growth hormone genes; 2) Use of arbitrary DNA primers (randomly amplified polymorphic DNA, or "RAPD" primers) to detect variation at large numbers of nuclear genes.

We used the polymerase chain reaction (PCR) to amplify variable sections (introns) of two growth hormone genes (GH-I and *GH-2*) in several salmonid species. Coho salmon had three DNA length variants for *GH-1* intron C. Restriction analysis and sequencing provided valuable information about the mode of evolution of these DNA sequences. We tested segregation of the variants in captive broods of coho salmon, and demonstrated that they are alleles at a single Mendelian locus.

Population studies using the *GH-1* alleles showed highly significant frequency differences between Lower Columbia River and Oregon Coast coho salmon, and marginal differences among stocks within these regions. These new markers are adequately defined and tested to use in coho salmon population studies of any size. The nature of the variation at *GH-1* (Variable Number Tandem Repeats, or "VNTRs") suggests that more genetic variants will be found in coho salmon from other areas.

GH-2 intron C also showed length variation in coho salmon, and this variation was found to be sex-linked. Because PCR methods require minute amounts of tissue, this discovery provides a technique to determine the gender of immature coho salmon without killing them. Chinook salmon had restriction patterns and sequence divergences similar to coho salmon. Thus, we expect that sex linkage of *GH-2* alleles predates the evolutionary divergence of Pacific salmon species, and that gender testing with this system will work on the entire group. Rainbow trout do not show this sex-linked variation.

Genetic markers detected by DNA amplification using arbitrary 10-basepair primers (Randomly Amplified Polymorphic DNA, or "RAPD" markers), are the newest and most promising method of assessing variation at large numbers of genetic loci. We have demonstrated the inheritance of these markers in rainbow trout, and we have found multiple variable genetic markers in coho salmon. Feasibility studies on the use of RAPDs on large salmon collections are described.

INTRODUCTION

Electrophoretic analysis of soluble enzymes and restriction enzyme analysis of mitochondrial DNA have become the mainstays of applied population genetics. As useful as these techniques are, they detect only a portion of the genetic variation that DNA sequencing shows to be present for most genes in natural populations. However, the application of newer DNA technologies to population genetic analysis, particularly for nuclear genes, has progressed slowly (Lewontin, 1992). Nevertheless, the prospect of increased allelic resolution and the opportunity to study classes of genes other than soluble enzymes are worthy reasons to pursue nuclear DNA methods.

Because of the greater development effort and higher costs involved, the primary focus must be on kinds of DNA variation that can be most efficiently detected. The actual number of alleles at the DNA level at any locus is rarely known, but it is expected to be greater than the number of electrophoretic alleles (Ramshaw et al., 1978). In a study designed to detect the full extent of variation at one locus, Kreitman (1983) found 43 DNA polymorphisms in a 2721-bp region of the *ADH* locus in *Drosophila melanogaster*, while allozyme analysis detected only the standard two (fast and slow) alleles. This work showed substantial variation in non-coding introns and flanking regions of the ADH locus.

Sequencing of multiple alleles is a feasible way to detect allelic variants, but a more efficient method than sequencing must be found to survey populations for this variation. Single nucleotide substitutions in PCR-amplified DNA can be detected by denaturing gradient gel electrophoresis (Lessa, 1992), single strand conformational polymorphisms (SSCPs, Orita et *al.*, 1989), or restriction enzyme digestion (Karl and Avise, 1992). Each of these techniques, however, detects only a portion of the base substitutions that differ among alleles; only sequencing can detect all such differences.

Much genetic variation in introns and other non-coding DNA is due to mutations that are larger and easier to detect than base substitutions. Insertions and deletions of multiple basepairs result in length variants that are easily detected on electrophoretic gels. Furthermore, many non-coding DNAs contain tandemly repeated sequences. Tandem repeats are prone to mutation by strand mispairing mechanisms that are particular to such sequences (Tautz *et al.*, 1986; Tautz, 1989; Harding *et al.*, 1992). Thus, multiple alleles are often found that differ in number of repeat units (e.g., micro-satellite sequences, mini-satellite sequences).

Length polymorphisms are sufficiently common in most genomes that they are an appropriate focus for population genetic analysis (Tautz, 1989). However, most DNA-based techniques potentially detect variation at several related loci. Multi-locus systems are not suitable for many population-genetic tests. Therefore, it is often necessary to show with inheritance data that DNA sequence variants segregate as alleles at a single locus.

Demonstrating allelism of DNA markers is particularly important in salmonid fishes because salmonids have many duplicated genes. The family Salmonidae has a tetraploid genome, due to a genome duplication event in the

lineage ancestral to the family. Disomic inheritance has re-evolved at many nuclear loci, but there remain two functional copies of many genes (Allendorf and Thorgaard, 1984; Johnson *et al.*, 1987). PCR primers for known genes designed without detailed knowledge of differences between duplicated loci will likely amplify sequences from both loci. The two growth hormone *(GH)* genes are the only duplicated pair of genes in salmonids for which both loci have been sequenced (Agellon *et al.*, 1988b). Knowledge of these sequences enabled us to design locus-specific primers for either *GH* gene.

This paper describes PCR amplification of introns of the two growth hormone genes in coho salmon *(Oncorhynchus kisutch)*. We began a search for genetic polymorphism in coho salmon with this approach because of the relative lack of isozyme variation in this species and the need for finding genetic markers capable of distinguishing between wild native and introduced hatchery coho salmon in the Lower Columbia River. We designed PCR primers located in coding DNA flanking intron C. An intron is expected to have relatively high genetic polymorphism, because its non-coding DNA can tolerate base changes and length variation without effects on fitness. We selected intron C because its size (540 bp in rainbow trout *GH-2;* Agellon et *al.*, 1988a) is convenient for direct sequencing of both strands. Here we report on a series of length variant alleles at growth hormone 1 *(GH-I)* and growth hormone 2 *(GH-2)* intron C in coho salmon. We also describe a feasibly study of RAPD (Randomly Amplified Polymorphic DNA) genetic markers (Williams *et al.*, 1990) in salmonid fishes.

MATERIALS AND METHODS

Pooulation Samples The following collections of coho salmon were provided by the Oregon Department of Fish and Wildlife, Portland: Sandy Hatchery, lower Columbia River, Oregon (N = 33, Stock 11, smolts, 5/11/92); Rock Creek Hatchery, Umpqua River, Oregon (N = 30, Stock 55, smolts, 4/28/92); Nehalem River wild 1, north coast Oregon (N = 28, presmolts, 6/24/92). Additional coho salmon were provided by the National Marine Fisheries Service, Seattle: Nehalem River wild 2, north coast Oregon (N = 33, fry, 5/14/92); Clackamas River wild 1, lower Columbia River (N = 35, NMFS #30747, fry, 5/15/92); Clackamas River (North Fork Dam) wild 2 (N = 19/20, NMFS #30746, numbered smolt tails, 5/15/92). Experimental broods of the Arlee strain of rainbow trout (0. *mykiss)* and the Anaconda broodstock of westslope cutthroat trout *(0. clarki lewisi)* are maintained by our laboratory. The Arlee strain of rainbow trout is founded from multiple wild stocks. It has high diversity for allozymes relative to natural populations (Leary *et al.*, 1983) and two very distinct mtDNA haplotypes (S.Forbes, unpublished data). The Anaconda broodstock of Westslope cutthroat trout is also derived from multiple sources. Three additional fry in the Clackamas Wild 1 collection were identified as chinook salmon by diagnostic allozyme alleles at *CK-A1*, *CK-A2* and *PEP-A1* (R. Leary, data not shown). In November 1992 we obtained coho salmon gametes from the Fall Creek Hatchery (Alsea River, mid-coast Oregon) to make experimental crosses. Tissues from the parents were kept for genetic analysis.

Methods DNA isolations were performed on 50-100mg muscle or liver tissue by the phenol/chloroform extraction method of Wirgin *et al.* (1990). PCR primer sequences for growth hormone intron C were located at positions in *GH-1* and *GH-2* exons 2 and 3 that are conserved among published rainbow trout, coho salmon, and chum salmon sequences (Agellon et *al.*, 1988b; Gonzalez-Villasenor *et al.*, 1988; Sekine *et al.*, 1985,1989). A single nucleotide site differs between the loci at the 3-prime end of each primer. These primers are perfect matches to the published sequences of rainbow trout *GH-1* and *GH-2* and coho salmon GH-1.

	Left	Right
GH-1	5'-ATCGTGAGCCCAATCGACAAGCAC-3'	5'-GGGTACTCCCAGGATTCAATCAGA-3'
GH-2	----------A----------G	----------G

Internal primers were also used on the longer *GH-1* introns in coho salmon: left 5'-GCAAGCAGACCACCAATTATG-3'; right 5'-AAACCACCGATTAGACAGTGATG-3'.

PCR reactions contained 100mM Tris-HCl pH 8.3, 500mM KCl, 4.0mM $MgCl_2$, 2ug/ml BSA, 0.2mM each dNTP, 0.5uM each primer, 0.5U Taq DNA polymerase (Perkin-Elmer Cetus or Promega Corp.), and 50-100ng template DNA in a total volume of 25ul. Thirty cycles of 92°C 1 min., 62°C 1 min., and 72°C 1 min. were followed by 3 min. at 72°C. The internal primers for *GH-1* used the same thermal program except with 56°C annealing for 1 min. Two to ten ul of each PCR reaction were run on 2.5% agarose gels in TAE buffer containing ethidium bromide. One ug 1Kb DNA ladder (Gibco-BRL) was used as a size standard.

Restriction digests of PCR products contained 2-10 ul of each reaction and 5 units of enzyme in a total volume of 20 ul, adjusted to the manufacturers' (Boehringer Mannheim Biochemicals) recommended conditions. Restriction products were run on agarose gels as above, or on 6% or 12% polyacrylamide gels in TBE, and stained with ethidium bromide. *GH-1* and *GH-2* PCR products were screened with 34 restriction endonucleases: *Alu*I, *Apa*I, *Ava*I, *Ava*II, *Asn*I, *Bam*HI, *Bcl*I, *Bgl*I, *Bgl*II, *Bst*EII, *Cfo*I, *Dde*I, *Dra*I, *Eco*RI, *Eco*RV, *Hae*III, *Hin*dII, *Hin*dIII, *Hin*fI, *Hpa*I, *Kpn*I, *Msp*I, *Nci*I, *Nde*II, *Nhe*I, *Pst*I, *Pvu*II, *Rsa*I, *Sca*I, *Scr*FI, *Stu*I, *Sty*I, *Taq*I, *Xba*I.

DNA was prepared for sequencing by running 20ul PCR reaction product in a 2.5% agarose gel and excising the desired bands. Single alleles for length variants were isolated as bands of unique size from heterozygous individuals. If necessary to produce enough template, single alleles were re-amplified from a gel slice soaked in 100 ul TE and frozen overnight at -40°C. Template DNAs were purified from gel slices using GeneClean (Bio101, Inc.). Automated sequencing was done by dye termination PCR cycle sequencing, using the original amplification primers, in an Applied Biosystems Incorporated 373A automated DNA sequencer. An internal sequencing primer was used for some templates: left primer - 5'-CCCCGATTGTCTAAACTCC-3'.

RESULTS

GH-1 and *GH-2* intron C PCR Products Two primer-specific size classes of GH intron C products corresponding to the *GH-1* and *GH-2* loci were produced in amplifications from each salmonid species tested (Table 1). *GH-1* products were about 820 bp, except in coho salmon, where they were several hundred bp longer. *GH-2* products were about 540 bp in all four species. This closely matched the size predicted from a genomic clone of rainbow trout *GH-2* (Agellon *et al.,* 1988b). Products from both genes showed length variation in coho salmon: three sizes at *GH-1* (Table 1; Fig. 1) and two sizes at *GH-2* (Table 1).

All *GH-1* and *GH-2* PCR products from all four species were digested with *Hinf*I (Table 1). All sites were conserved among species and among alleles except one distinguishing the *GH-2* products in coho salmon. We also digested apparent homozygotes for each allele at *GH-1* and *GH-2* in coho salmon with 33 other restriction enzymes (Table 2). Eighteen enzymes cut one or more alleles. At *GH-1,* there were no restriction site differences among length variants for any enzyme. In contrast, the two alleles at *GH-2* differed at six restriction sites. Restriction digestion did not reveal more variants at either locus than were detectable by PCR product length analysis alone.

Automated cycle sequencing provided accurate, reproducible sequence data to at least 300 bp from either primer. Opposite strand sequences usually overlapped in the middle, except for coho *GH-1* sequences, which were too long to sequence fully from the ends. Internal primers for coho *GH-1* gave sequences that read through the central tandem repeat region. Figure 3 shows the longest allele (*GH-1*a*), with a consensus sequence of the tandem repeat region in all three alleles. The number of repeats in each of the three alleles explains the length differences seen on agarose gels. The AGCC motif that punctuates the 31-bp, AT-rich repeats gave a strong signal throughout. However, because sequencing accuracy was poor in this region, probably because of its repeated structure, it is not certain that all repeat units are identical. *GH-2* from rainbow trout nearly matched the published sequence for a genomic clone of a *GH* gene (Agellon et *al.,* 1988b). Positions of all restriction sites shown in Table 2 were later confirmed by sequencing.

Interspecies sequence divergences of this *GH-1* intron are higher than for *GH-1* exons, and lower than for mitochondrial DNA. Evolutionary divergences among alleles and among species for *GH-1* sequences were measured by aligning al1 five sequences, discounting the gaps, and counting the number of mismatched bases (Table 3). Exon sequences (cDNAs) for this comparison were available only for coho salmon (Gonzales-Villasenor et *al* ., 1988) and rainbow trout (Agellon *et al.,* 1988b). The mean divergence of intron C between the three coho alleles and rainbow *GH-1* is 2.5 percent (Table 3). This is nearly twice the value (1.4%) for *GH-1* exon sequences. In a 2214-bp segment of mitochondrial DNA the coho-rainbow divergence is 7.2 percent (Thomas and Beckenbach, 1989).

We tested segregation of the coho salmon length variants at *GH-1* in experimental crosses (Table 4; Fig. 4). When the parents were alternate homozygotes, all 20 progeny were heterozygotes. When either parent alone was heterozygous, the 20 progeny showed 1:l segregation from that parent. The

length variants produced by *GH-1* intron C primers in coho salmon are clearly allelic.

GH-1 genotypes were recorded for 203 coho salmon in seven wild and hatchery collections (Table 5). All samples conform to expected binomial (Hardy-Weinberg) genotype proportions, suggesting that they are from random-mating populations of reasonable size. We tested all pairs of collections for gene frequency differences with 2 x 3 contingency tests (Table 6). The seven collections fall clearly into a coastal Oregon group and a Lower Columbia River group, since most pairwise tests between these clusters fall below the p=0.001 significance level (Table 6, Fig. 5). The only exception is Rock Creek Hatchery (Umpqua River, Oregon), which more closely resembles the lower Columbia stocks than the coastal stocks. Wild fish in each group do not differ significantly from the hatchery in that group (Sandy Hatchery vs. Clackamas wild, Fall Creek Hatchery vs. Nehalem wild).

Sex linkage of *GH-2* We designated the two coho *GH-2* alleles *GH-2*a* (530 bp) and *GH-2*b* (551 bp). The 21-bp difference is detectable on high concentration agarose gels. However, we routinely identified *GH-2*b* by a diagnostic 456-bp HinfI restriction fragment (Table 1; Fig. 2; Fig. 6). In collections of coho salmon large enough to sex (>100mm), all the males had a bright GH-2*b band, while all the females had no *GH-2*b*. The association of *GH-2*b* and phenotypic male sex was perfect in a total of 27 female and 36 male coho salmon (Table 7). The frequency of the male-specific marker in collections of fry that were too small to sex was 0.52 in Nehalem River wild coho and 0.51 in Clackamas River wild coho, consistent with sex linkage in these populations as well (Table 7). Progeny in three experimental crosses had the paternal band in one-half the progeny (Table 8; Fig. 2).

We also studied restriction patterns and sequences of *GH-2* markers in other salmonid species. *GH-2* restriction patterns from chinook salmon appear similar to the coho salmon patterns. That is, there are two types of individuals that correspond to the *aa female and **ab* males seen in coho salmon on the basis of the size of the products and the HinfI restriction patterns (Table 1). Rainbow trout and westslope cutthroat trout (subgenus Parasalmo) have less *GH-2* variation than the two salmon species. There was no length variation in *GH-2* products among 71 rainbow trout and 22 westslope cutthroat trout. There was also no HinfI restriction site variation at *GH-2* in 69 rainbow trout and 12 cutthroat trout.

DNA sequences of the *GH-2* alleles from coho salmon (Fig. 6) corroborated the striking degree of divergence indicated by restriction digests. All *GH-2*a* sequences were putative diploids because the allele is X-linked; *GH-2*b* sequences are of single alleles. We compared pairs of sequences by aligning the sequences, discounting the gaps, and counting the DNA base mismatches (percent sequence divergence; Table 9). The sequence divergence between the *GH-2*a* and *GH-2*b* alleles is remarkable. In coho salmon, *GH-2*a* and *GH-2*b* differ by 6.5 percent, and in chinook salmon these alleles differ by 7.4 percent (Table 9). These intraspecies differences are larger than the divergence of either allelic type between species (0.9 percent for *GH-2*a*, 3.9 percent for *GH-2*b*; Fig. 7).

Another difference between *GH-2*a* and *GH-2*b* was in the relative intensity of the amplified bands. *GH-2*b* was always bright when present, but *GH-2*a* varied from a brightness equal to *GH-2*b* to near invisibility. A faint band matching *GH-2*a* in size appeared in *GH-1* amplification from every coho individual (Fig. 1). Restriction enzymes which cut *GH-2*a* also cut this band. Sequencing of the band confirmed it to be identical to *GH-2*a* (data not shown). However, no *GH-2*b* product was seen in any *GH-1* amplification. We infer from this that *GH-2*a* alleles have a cryptic change in a PCR priming sites that matches the 3' end of a *GH-1* primer. This mutation results in weak amplification with *GH-1* primers. Also, primer site polymorphisms apparently cause variable amplification efficiency of *GH-2*a* relative to *GH-2*b* among individuals. We took advantage of weak *GH-2*a* amplification in selected individuals to obtain clean sequencing templates for *GH-2*b*.

Nuclear DNA "RAPD" Markers We also tested the RAPD (Randomly Amplified Polymorphic DNA) technique for measuring genetic differences among populations. These genetic markers detected by DNA amplification using arbitrary 10-basepair primers are a new and promising method of assessing variation at large numbers of genetic loci (Williams et *al.*, 1990). However, the technique is so new that there are no published accounts of use of the method in salmonid fishes, and none discussing the use of such data for population genetic purposes. We have demonstrated the inheritance of 14 of these markers in the Arlee strain of rainbow trout (data not shown). We have also found 14 variable genetic markers in coho salmon DNAs using 15 different RAPD primers (Table 10). Since there are over 500 RAPD primers commercially available, this method can potentially screen a vast number of genetic markers in the coho salmon nuclear genome. A test for marker frequency differences between Sandy and Rock Creek hatchery coho using six variable RAPD markers revealed a marginally significant ($P<0.1$) difference (Table 11). This is concordant with the analysis of these collections with the *GH-1* markers. Efforts to extend this analysis to more primers and more collections met with technical difficulties (see Discussion).

DISCUSSION

Choice of Molecular Genetic Techniques Appropriate choice of a molecular technique for studying population structure depends upon: (1) The fundamental ability of a technique to reveal genetic variation, (2) The nature of the variation actually present in the study populations, (3) The cost-effectiveness of the technique for surveying large numbers of loci in many individuals, and (4) The facility with which a general strategy, once developed, can be applied to new loci or to new study species.

Although allozyme analysis can examine large numbers of loci, it reveals only a portion of the total allelic variation at each locus, and it only addresses a single class of structural genes. Mitochondrial DNA restriction analysis or sequencing often reveals a larger number of variants. Fourteen haplotypes were found in Washington State coho populations (P. Moran, personal communication), and seven haplotypes were found in pilot studies of Lower Columbia River coho salmon (K. Currens, personal communication). However,

mtDNA is effectively only one locus, and its haploid, maternal mode of inheritance makes it not necessarily representative of the nuclear genome.

Allozymes are still frequently the method of choice for many fisheries applications (Utter 1991). However, some species have too little allozyme variation to be useful. Lower Columbia River coho salmon fall in this category. Variant allozyme alleles in coho are generally rare (Wehrhahn and Powell, 1987; Reisenbichler and Phelps, 1987; Johnson et *al.*, 1991) giving little statistical power to study population differences.

PCR Amplification of Nuclear Introns We have found that genetic variation that is detectable as PCR product length differences can be efficiently surveyed in natural populations. This result is far more important than the fact that growth hormone genes in particular are useful for this purpose. Whether length variation is common in other salmonid genes is not yet known. We expect that such variation is quite gene- and species-specific, since coho salmon showed length variants in this study, and rainbow trout and cutthroat trout did not. Nevertheless, many non-coding DNAs in all species display length variation, and this study shows that length variants can be efficiently scored, once detected. We estimate the cost of using a developed single-locus system such as *GH-1,* at about $1.50 per fish. Since much of this cost is for preparation of sample DNAs, the cost would be less for additional loci. This cost does not include labor, development or equipment costs.

The polymerase chain reaction is a boon to population geneticists. It permits genetic studies to be done using minuscule, non-lethal samples, and it has been found very effective for mitochondrial DNA studies in fishes (Beckenbach, 1991; Whitmore et *al.*, 1992). It also potentially targets genetic variation that is not accessible to other methods. However, as the present study shows, development of a system to detect variation in known diploid nuclear genes, especially in a species with many duplicated genes, is not simple. Because locus-specific primers must be devised and the inheritance of markers tested, the promise of this approach for salmonid population genetics will improve only as fast as the database of single copy gene sequences in salmonid fishes grows.

Population Structure of Coho Salmon *GH-1* allele frequencies show a striking difference between Oregon coast and Lower Columbia River (LCR) coho salmon stocks. Genetic distinctiveness is one of several factors involved in identifying Evolutionarily Significant Units (ESUs) under the Endangered Species Act (Waples, 1991). Our results provide evidence that Columbia River and coastal coho salmon belong in separate ESUs.

Another important comparison is that of putative wild coho salmon and local hatchery stocks. Evidence of a genetically unique wild-spawning coho salmon run in the Lower Columbia River would warrant its listing as an ESU (Oregon Trout and Pacific Power, 1991; Watkins and Vigg, 1991). We found no difference in *GH-1* allele frequencies between two samples of putative wild fish from the Clackamas River and one collection from Sandy Hatchery, which has had historical contributions from the Clackamas as well as from other lower Columbia hatcheries. Neither were the Fall Creek hatchery coho salmon

significantly different from wild Nehalem River coho salmon. We must emphasize, however, that although a gene frequency difference can demonstrate that two populations do not randomly interbreed, a lack of differentiation at one locus does not prove two stocks are not different. Stock structure analysis with *GH-1* markers can be enhanced in the future by testing larger collections of fish, but it will be even better enhanced by looking at more different genes (Nei, 1987).

GH-2 alleles are sex-linked The *GH-2*b* allele detected by *Hinf*I digestion of *GH-2* specific PCR products is unequivocally Y-chromosome linked in coho salmon. We expect that sex linkage of *GH-2* alleles predates the evolutionary divergence of Pacific salmon species, because of the apparent homologies between coho and chinook salmon for *GH-2*a-* and *GH-2*b-* type alleles.

Because PCR methods require minute amounts of tissue, this discovery provides a technique to determine the gender of immature salmon from a small fin clip without killing the fish. Such a procedure was developed using a different Y-linked DNA marker in chinook salmon (Devlin et *al.*, 1991), but it works only on chinook salmon. Because the *GH-2* alleles are conserved in coho and chinook salmon, we expect that they will work for sex identification in other Pacific salmon species not yet tested. Rainbow trout and cutthroat trout, however, do not show sex-specific length variation in this *GH-2* intron. Future comparative studies of this region of the genome in trout and salmon will be highly informative about the evolution of sex determination.

RAPD Markers The use of RAPDs for population genetics is unproved. The strength of the RAPD technique lies in the opportunity to assess a large number of markers within a single controlled experiment. The main weakness is a technical sensitivity that makes it difficult to replicate results. The coho results described above are a best-case example of controlled conditions. All the DNAs were isolated by the same procedure, and reactions were run consecutively on a single PCR machine under identical conditions. The results appear reliable. However, mechanical difficulties with PCR machines have prevented us from replicating these results, or extending the study. Another lab could use the same primers with the same fish, and likely would get a similar measure of genetic distance, but their data and ours would not be commensurable on a marker-for-marker basis.

We do not feel that the RAPD technique as it exists is robust enough to serve widely in salmonid stock identification, where comparison of data among different collections from different labs is necessary. At present, even within one lab, the need for replicate controls to assure comparability among data sets seriously compromises the theoretical advantages of the technique. However, RAPDS have greater promise for focal studies of pairs of populations of particular interest, where broader comparisons are not necessary. This prospect warrants research on technical improvements to make the RAPD technique more consistent. For genetic mapping studies, where tests within single experiments are highly informative, RAPDs are already promising. Our initial studies showed good Mendelian inheritance of RAPD markers within families of rainbow trout. We have also begun a study to gene-centromere map variable RAPD bands within gynogenetic diploid broods of rainbow trout.

LITERATURE CITED

Agellon, L.B., S.L Davies, T.T. Chen, and D.A. Powers. 1988a. Structure of a fish (rainbow trout) growth hormone gene and its evolutionary implications. Proc. Natl. Acad. Sci. U.S.A. 85:5136-5140.

Agellon, L.B., S.L Davies, C.-M. Lin, T.T. Chen, and D.A. Powers. 1988b. Rainbow trout has two genes for growth hormone. Mol. Reprod. Dev. 1:11-17.

Allendorf, F.W. and G.H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes, pp. 1-53. In B. Turner (ed.), Evolutionary Genetics of Fishes. Plenum Press, New York.

Beckenbach, A.T. 1991. Rapid mtDNA sequence analysis of fish populations using the polymerase chain reaction (PCR). Canad. J. Fish. Aquat. Sci. 48:95-98.

Devlin, R.H., B.K. McNeil, T.D.D. Groves, and E.M. Donaldson. 1991. Isolation of a Y-Chromosomal DNA probe capable of determining genetic sex in chinook salmon (Oncorhynchus *tshawytscha)*. Can. J. Fish. Aquat. Sci. 48:1606-1612.

Gonzalez-Villasenor, L.I., Zhang, P., Chen, T.T. and D.A. Powers. 1988. Molecular cloning and sequencing of coho salmon growth hormone. Gene 65:239-246.

Hardinq, R.M., A.J. Boyce, and J.B. Cleqq. 1992. The evolution of tandemly repetitive DNA: recombination rules. Genetics 132:847-859.

Johnson, K.R.; Wright Jr., J.E.; May, B. 1987. Linkage relationships reflecting ancestral tetraploidy in salmonid fish. Genetics 116:579-591.

Johnson, O.W., T.A. Flagg, D. J. Maynard, and F.W. Waknitz. 1991. Status Review for lower Columbia River Coho Salmon. U.S. Dep. Commerce, NOAA Technical Memorandum. 96 pp.

Karl, S.A., and J.C. Avise. 1992. Balancing selection at allozyme loci in oysters - Implications from nuclear RFLPs. Science 256:100-102.

Kreitman, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature 304:412-417.

Leary, R.F., F.W. Allendorf and K.L. Knudsen. 1983. Developmental stability and enzyme heterozygosity in rainbow trout. Nature 301:71-72.

Lessa, E. P. 1992. Rapid surveying of DNA sequence variation in natural populations. Mol. Biol. Evol. 9:323-330.

Lewontin, R.C. 1992. Electrophoresis in the development of evolutionary genetics: milestone or millstone? Genetics 128:657-662.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Oregon Trout and Pacific Power. 1991. Request to the National Marine Fisheries Service to reconsider its decision on the wild Lower Columbia River coho under the Endangered Species Act. Oregon trout, Portland, Oregon. Pacific Power, Portland, Oregon. 21 pp.

Orita, M., I. Hiroyuki, H. Kanazawa, K. Hayashi, and T. Sekiya. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single strand conformation polymorphisms. Proc. Nat. Acad. Sci. USA 86:2766-2770.

Ramshaw, J.A.M., J.A. Coyne, and R.C. Lewontin. 1979. The sensitivity of gel electrophoresis as a detector of genetic variation. Genetics 93:1019-1037.

Reisenbichler, R.R., and S.R. Phelps. 1987. Genetic variation in chinook, *Oncorhynchus tshawytscha,* and coho, *0. kisutch*, salmon from the north coast of Washington. Fish. Bull. 85:681-701.

Sekine, S., T. Mizukami, T. Nishi, Y. Kuwana, A. Saito, M. Sato, S. Itoh, and H. Kawauchi. 1985. Cloning and expression of cDNA for salmon growth hormone in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 82:4306-4310.

Sekine, S., T. Mizukami, A. Saito, H. Kawauchi, and S. Itoh. 1989. Isolation and characterization of a novel growth hormone cDNA from chum salmon *(Oncorhynchus keta)*. Biochim. Biophys. Acta 1009:117-120.

Tautz, D., M. Trick, and G.A. Dover. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nuc. Acids Res. 17:6463-6471.

Thomas, W.K., and A.T. Beckenbach. 1989. Variation in salmonid mitochondrial DNA: Evolutionary constraints and mechanisms of substitution. J. Mol. Evol. 29:233-245.

Utter, F.M. 1991. Biochemical genetics and fishery management - An historical perspective. J.Fish Biol. 39:1-20.

Waples, R.S. 1991. Definition of "Species" Under the Endangered Species Act: Application to Pacific Salmon. U.S. Dep. Commerce, NOAA Technical Memorandum NMFS F/NWC-194, 29 pp.

Watkins, D. and S. Vigg. 1991. Comments of the Bonneville Power Administration on the National Marine Fisheries Service determination to not propose listing of the Lower Columbia River coho salmon under the Endangered Species Act. Portland, Oregon. 40 pp.

Wehrhahn, C.F., and R. Powell. 1987. Electrophoretic variation, regional differences, and gene flow in the coho salmon *(Oncorhynchus kisutch)* of southern British Columbia. Can. J. Fish. Aquat. Sci. 44:822-831.

Whitmore, D.H., T.H. Thai, and C.M. Craft. 1992. Gene amplification permits minimally invasive analysis of fish mitochondrial DNA Transactions of the American Fisheries Society 121:170-177.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids. Res. 18:6531-6535.

Wirgin, I.I., M. D'Amore, C. Grunwald, A. Goldman, and S.J. Garte. 1990. Genetic diversity at an oncogene locus and in mitochondrial DNA between populations of cancer-prone Atlantic tomcod. Biochemical Genetics 28:459-476.

Table 1. Fragment sizes of intron C PCR products from growth hormone 1 and growth hormone 2 in coho salmon, chinook salmon and rainbow trout. All restriction sites were confirmed by sequencing. Fragments less than 50 bp were detected by sequencing only. N = number of alleles tested, assuming diploidy at both loci.

Species	Locus/Allele	N	Product Size (bp)	Hinf I Fragments (bp) (5' - 3' order)					
Coho	*GH-l*a*	**153**	**1178**	27	134	631	34	334	**16**
Coho	*GH-l*b*	47	1111	27	134	566	34	334	**16**
Coho	*GH-1*c*	16	1083	27	134	532	34	334	**16**
Chinook	*GH-1*	6	825	27	134	280	36	332	**16**
Rainbow	*GH-1*	80	822	27	118	284	36	340	**16**
Coho	*GH-2*a*	204	530	27	--		488		**16**
Coho	*GH-2*b*	101	551	27	**52**		456		**16**
Chinook	*GH-2*a*	4	540	27	--		497		16
Chinook	*GH-2*b*	2	551	27	52		456		16
Rainbow	*GH-2*	**138**	540	27	53		444		**16**
Cutthroat	*GH-2*	**12**	540	27	53		444		**16**

Table 2. Restriction analysis of *GH-1* and *GH-2* intron C PCR products from coho salmon. 0 = enzyme does not cut allele. N = number of alleles tested. For *GH-1* only the **a* allele (1176 bp) is shown; the **b* and **c* alleles have the same restriction sites, and differ from **a* only in length.

	*GH-1*a*		*GH-2*a*		*GH-2*b*	
Enzyme	**bp**	**(N)**	**bp**	**(N)**	**bp**	**(N)**
Uncut	1176		530		551	
AluI	815/244/73/45	(18)	485/45	(4)	506/45	(2)
AsnI	0	(19)	0	(26)	460/91	(30)
AvaI	726/453	(14)	0	(6)	0	(3)
AvaII	0	(14)	477/53	(4)	498/53	(6)
DdeI	548/286/154/132/56	(16)	233/227/56/14	(54)	242/168/127/14	(38)
DraI	0	(14)	290/240	(16)	299/252	(8)
HaeIII	873/303	(32)	0	(16)	0	(7)
HindIII	0	(13)	0	(8)	505/46	(4)
HinfI	631/334/134/34/27/16	(153)	487/27/16	(204)	456/52/27/16	(101)
MspI	875/301	(13)	0	(6)	0	(4)
NciI	875/301	(13)	0	(6)	0	(4)
NdeII	0	(13)	389/141	(4)	398/153	(6)
PstI	777/399	(13)	0	(4)	0	(2)
PvuII	814/362	(13)	0	(4)	0	(2)
RsaI	1002/114/60	(14)	248/167/115	(6)	260/178/114	(3)
ScrFI	823/302/51	(13)	479/51	(3)	500/51	(2)
StyI	0	(13)	369/161	(4)	378/173	(2)
TaqI	526/385/201/64	(13)	0	(14)	434/118	(12)

Table 3. *GH-1* **intron C percent DNA sequence divergence. The analysis used pairwise divergences at 702 total nucleotide sites aligned among the five sequences.**

Species/Allele		Coho *GH-l*a*	Coho *GH-l*b*	Coho *GH-1*c*	Rainbow *GH-1*	Chinook *GH-1*
Coho	*GH-l*a*	-				
Coho	*GH-l*b*	1.1	-			
Coho	*GH-l*c*	1.4	0.7			
Rainbow	*GH-1*	3.0	2.3	2.3	-	
Chinook	*GH-1*	5.5	4.8	4.8	3.6	-

Table 4. Segregation of *GH-1* **intron C PCR product length variants** in captive broods of coho salmon. *GH-l*a* = **1178 bp;** *GH-l*b* = **llll bp.**

Family	Parental Genotypes		Offspring Counts		
	Female	Male	*aa*	*ab*	*bb*
u2	*ab*	*bb*	0	10	10
u3	*ad*	*bb*	0	20	0
u7	*da*	*ab*	8	12	0

Table 5. Genotype and allele frequencies at the *GH-1* locus in hatchery and wild collections of coho salmon. To test for expected binomial genotype proportions the less common alleles (*b and *c) were pooled. ns = not significant.

		GH-1 Genotype Counts						Allele Freq.			
Collection	N	*aa*	*ab*	*bb*	*ac*	*bc*	*cc*	*a*	*b*	C	X^2
Sandy Hatchery	33	26	4	0	3	0	0	0.894	0.061	0.045	0.49^{ns}
Clackamas Wild 1	34	18	9	0	6	0	1	0.750	0.132	0.118	1.05^{ns}
Clackamas Wild 2	20	13	3	0	3	0	0	0.842	0.079	0.079	0.90^{ns}
Rock Cr. Hatchery	30	15	7	0	7	1	0	0.733	0.133	0.133	1.12^{ns}
Fall Cr. Hatchery	25	6	11	6	1	1	0	0.480	0.480	0.040	0.04^{ns}
Nehalem Wild 1	28	8	12	4	1	3	0	0.518	0.411	0.071	0.38^{ns}
Nehalem Wild 2	33	15	10	5	3	0	0	0.652	0.303	0.045	0.58^{ns}

Table 6. Tests for *GH-1* allele frequency differences between collections. Data are from Table 4. All tests are 2 x 3 chi-square contingency tests (df=2). Significances are determined by Monte Carlo simulation. *<P0.05, **P<0.01, ***P<0.00l.

	Clackamas 1	Clackamas 2	Rock Cr.	Fall Cr.	Nehalem 1	Nehalem 2
Sandy Hatchery	4.75	0.52	5.52	27.56***	23.08***	13.18***
Clackamas Wild 1	-	1.52	0.07	17.81***	12.46**	7.10
Clackamas Wild 2			1.90	17.36***	13.64***	7.70*
Rock Cr. Hatchery				16.71***	11.55**	7.16*
Fall Cr. Hatchery					0.82	3.82
Nehalem Wild 1						2.23

Table 7. Association of genotypes of coho salmon at *GH-2* with gender. The **a* and **b* alleles were scored by the size of *HinF*I restriction fragments (Fig. 2).

Collection	Females	Males	Unknown		Freq.
	aa	*ab*	*ad*	*ab*	*GH-2*ab*
Sandy Hatchery	14	19	--	--	0.58
Rock Cr. Hatchery	13	17	--	--	0.57
Fall Cr. Hatchery	20	5*	--	--	---
Nehalem Wild	--	--	32	29	0.52
Clackamas Wild	--	--	28	27	0.51
Total	47	41	60	56	0.52

***Fall Creek fish were selected on the basis of gender for experimental matings; therefore, the frequency of *GH-2*b* is not a random sample of this population.**

Table 8. Inheritance of variability at *GH-2* in experimental captive broods (24) of coho salmon. The **a* and **b* alleles were scored by the size of *HinF*I restriction fragments (Fig. 2).

Family	Parents		Offspring	
	Female	Male	*ad*	*ab*
u2	*aa*	*ab*	13	7
u3	*aa*	*ab*	10	10
u7	*aa*	*ab*	8	12

Table 9. *GH-2* **intron C percent sequence divergence. The analysis used pairwise divergences at 432 total nucleotide sites aligned among the four salmon sequences (above diagonal), and 419 sites aligned among all six sequences (below diagonal).**

		Chinook *GH-2*a*	Chinook *GH-2*b*	Coho *GH-2*a*	Coho *GH-2*b*	Rainbow *GH-2*	Cutthroat *GH-2*
Chinook	*GH-2*a*	-	7.4	0.9	6.0	-	-
Chinook	*GH-2*b*	*7.2*	-	*7.9*	*3.9*	-	-
Coho	*GH-2*a*	1.0	7.6	-	6.5	-	-
Coho	*GH-2*b*	*5.7*	4.5	6.2	-		
Rainbow	*GH-2*	**2.1**	**6.9**	2.6	5.0	-	-
Cutthroat	*GH-2*	3.6	7.9	4.1	6.4	1.9	-

Table 10. RAPD polymorphisms in hatchery coho salmon and hatchery rainbow trout.

	Coho	Rainbow
	--em--	
No. Fish Tested	10	19
No. Primers	15	15
No. Scorable Bands	51	49
No. Variable Bands	14	24

Table 11. Marker frequencies for RAPD primers A2 and A10 in two hatchery collections of coho salmon.

Marker	Sandy	Rock Cr.	Prob.
A2/1	0.83	0.87	1.00
A10/1	0.97	0.85	0.17
A10/2	0.14	0.00	0.12
A10/3	0.91	0.89	1.00
A10/4	0.25	0.44	0.17
A10/5	0.00	0.13	0.09
Combined Prob.			co.10

Figure 1. Polymorphism of *GH-1* **intron C PCR products in coho salmon. Lanes** 1-19: **three alleles** (1178 **bp,** 1111 **bp, 1083 bp) segregate at** *GH-1* **in** 19 **unrelated individuals. Lane 20:** 1Kb **DNA Ladder (Gibco-BRL).**

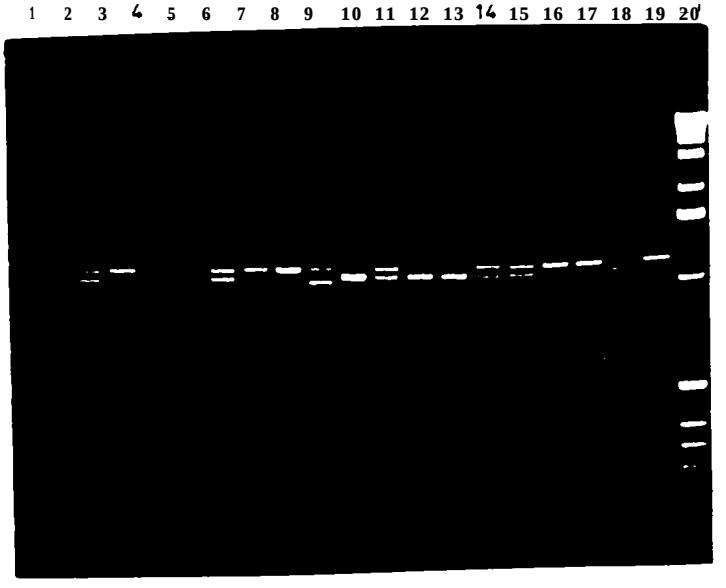

Figure 2. Segregation of *GH-2* alleles in coho salmon experimental brood U2. Lanes 1 and 18: 1Kb DNA Ladder (Gibco-BRL). *HinfI* digests of *GH-2* intron C PCR products are shown for the female and male parents (lanes 2 and 3) and 14 offspring (lanes 4-17), in a 2.5% agarose gel. The 488-bp fragment (lane 2) represents the *GH-2*a* product (530 bp total). The 456-bp fragment (lane 3) is characteristic of the Y-linked *GH-2*b* product (551 bp total). The brightness of *GH-2*a* in males varied markedly among unrelated individuals, but the *Y*-linked *GH-2*b* band was always bright when present. This variation is likely due to PCR primer site differences among *GH-2*a*-type alleles.

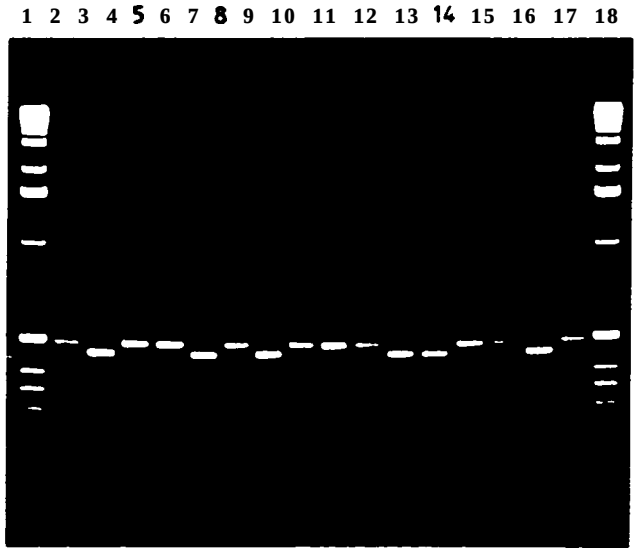

Figure 3. DNA sequences of coho salmon *GH-1* intron C PCR products. Nucleotides which are alignable with coho *GH-2* are marked (*). Three alleles (*GH-1*a,*b,*c*) have 11, 9 and 8 contiguous copies of the 31-bp tandem repeat, respectively. The entire *GH-1* sequence is alignable with rainbow trout and chinook salmon *GH-1* except for the coho tandem repeats, which are absent in these other species.

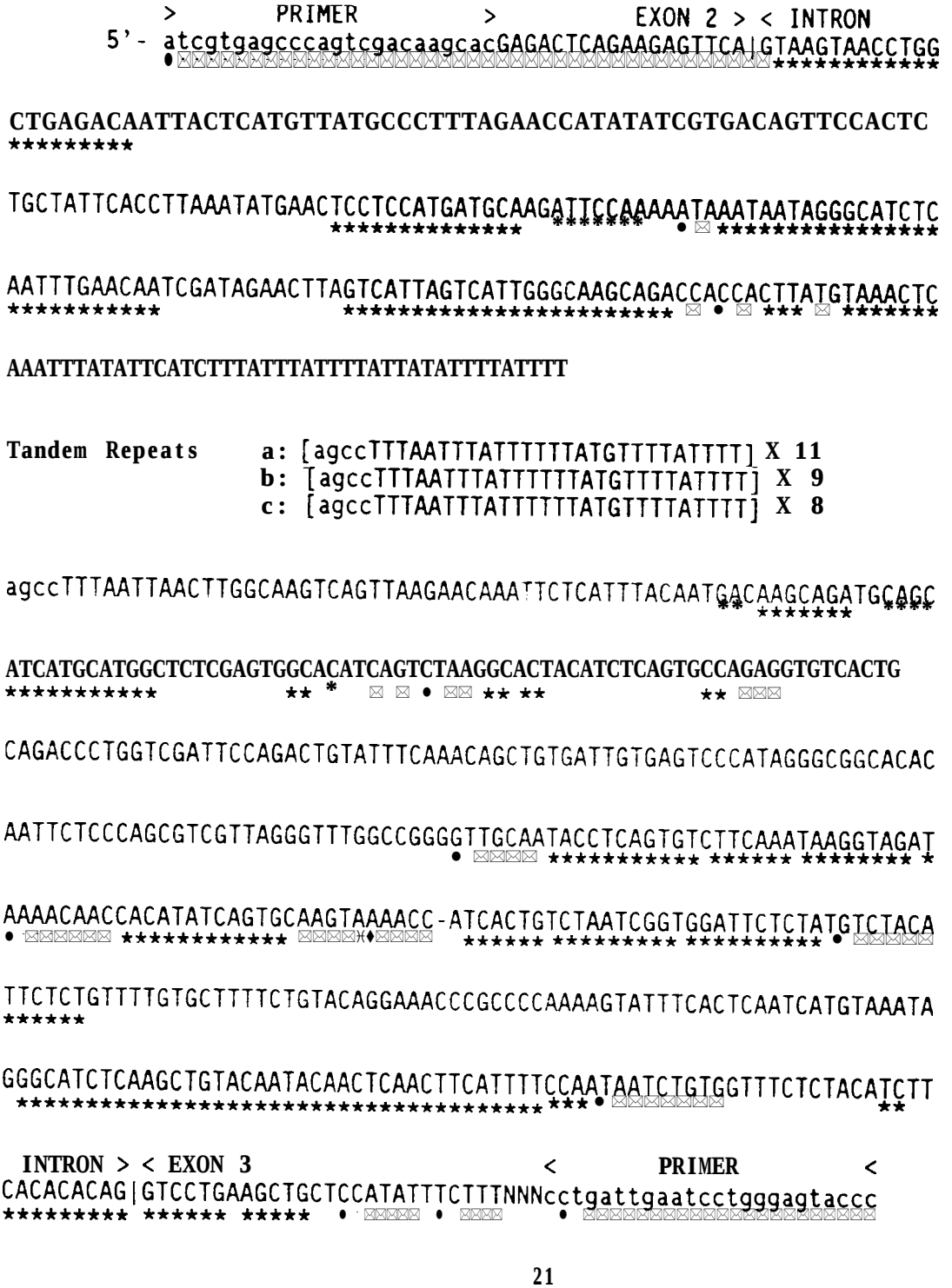

Figure 4. Segregation of *GH-1* alleles **a* and **b* in an experimental brood of coho salmon. Lanes 1 and 18: 1Kb DNA Ladder (Gibco-BRL). Lanes 2-4: unrelated adults. Lanes 5 and 6: female and male parents. Lanes 7-17: offspring.

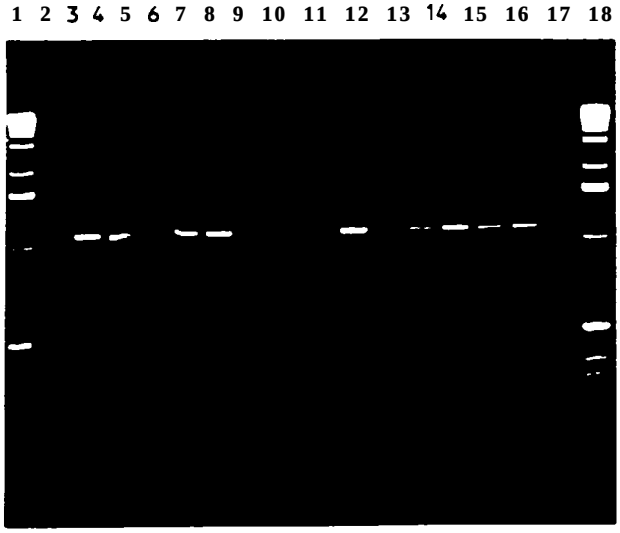

Figure 5. Dendrogram of genetic distances based on *GH-1* allele Frequencies in coho salmon collections. Data are from Table 5. Only differences between the Lower Columbia and Oregon Coast clusters are statistically significant (Table 6).

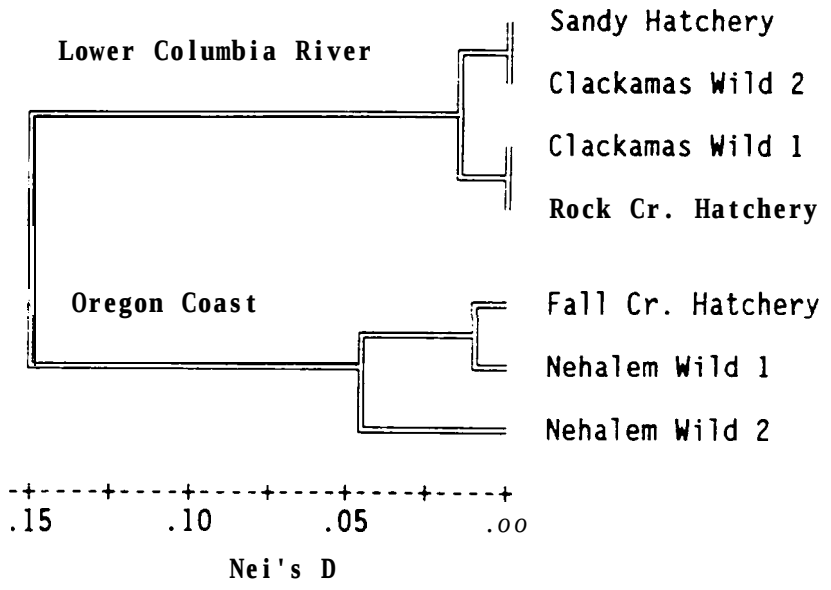

Figure 6. DNA sequence of coho salmon *GH-2* intron C PCR products. HinfI restriction sites are underlined, including the one distinguishing the **b* allele from the *a allele (79 bp). Nucleotide mismatches and alignment gaps are marked (*).

```
                >       PRIMER          > HinfI     EXON 3> <INTRON C
GH-2*a          atcgtgagcccaatcgacaagcagNNGACTCAGAAGAGTTCA|GTAAG  7
GH-2*b          atcgtgagcccaatcgacaagcagNNGACTCAGAAGAGTTCA|GTAAG  7

                                  HinfI
TTACCTGGCTGAGACAATCCTCCATGATGCACAATTCCAATAGGGCATCTC    107
TTACCTGGCGGAGACAATCC-GCACGATTGCACGATTCCAATAGGGCATCTC     6
         *          **  *     *    ⊠  •

AATTTGAACAA------------GTCATTAGTTATTGGGCAAGCAGATCCCCGATTGTC 154
AATTTGAATAATCGATACAACTTAGTCATTAGTTATTGGGCAAGCAGATCCCCAATTGTC 166
        *  **************                              *

TAAACTCCATGGGTAAATATATACTGTAGATAAGCAGAACCAGCATCATGCATGGTGG~ 214
TAAACTCCATGGGTAAATATATACTGTAGAAAAGCAGAACCAGCATCATGCATGGTGGAA 226
                          *

ATTAAATCTAGCCATGACAGGAAGTTTTAAATTGTACACTTAAAATCAACAGTAAAATGT 274
ATTAAATCTAGCCATGATAGGGAGTTTTAAATTGTACACTTAAAATCGGCAGTAAAATGT 286
                 *   *                          **

TGCTATACCTCAGTGCCTTCAACTAAGGTAGGTGAAAACATCACATATCACAGTCCTTG 334
TGCTATACCTCAGTGCCTTCAATTAAGGTAGGTTAAAACAACCACACACCATAGGCCTTG 345
                        •          ⊠       •    ⊠⊠  *  ⊠

TAAGTAAAACCCATCACTCTCTAATCGGCGGTTTCTCTACGTCTACATTCTCCAGCAATG 394
GAAGTAAAACTCATCACTCTCTAATCGGCGGGTTCTCTACGTCTACATTCTCCAGCCATG 405
•         •                  ⊠                          ⊠

TTT-ATGT---------GGCATCTCAAGCTGTACAATTACAACTCAACTTCATTTTCTAA 444
TATCATGTAAATGATATGTCATCTCAAGCTGTACAA-TACAATTCAACTTCATTTTCTAA 464
 • *    ********* *                  ⊠       ⊠

TCATCTGTGGTTTCTCTACATCTACACACACAG|GTCCTGAAGCTGCTCCATATCTCTTT 503
TAATCTGTGTTTTCTCTACATCTACACACACAG|GTCCTGAAGCTTCTCCATATCNNNNN 523
 *         *                 INTRON C> <EXON 4     *

   <       HinfI   PRIMER  <
CNNcctgattgaatcctgggagtaccc 530
NNNcctgattgaatcctgggagtaccc 551
```

Figure 7. UPGMA (unweighted pair group using arithmetic averages) cluster analysis of sequence divergence between *GH-2* intron C alleles in coho salmon and chinook salmon. The analysis used pair-wise divergences at 432 total nucleotide sites aligned among the four sequences (Table 9).